

郭亮. 基于 CGAN 的近红外关联成像高分辨率重构[J]. 智能计算机与应用, 2024, 14(5): 144-149. DOI: 10.20169/j.issn.2095-2163.240519

# 基于 CGAN 的近红外关联成像高分辨率重构

郭亮

(上海工程技术大学 保卫处, 上海 201620)

**摘要:** 由于复杂天气情况和夜间可视度低等场景限制, 红外图像技术应用广泛, 但是获取的成本较高。为降低红外图像的获取成本, 利用二维傅里叶单像素成像的方法获得低采样率的红外图像, 再利用条件对抗生成式网络进一步重构出清晰的红外图像。实验结果表明, 在采样率为 6.25% 的情况下, 可以重构出接近原始清晰图像的红外图像, 因而利用关联成像降低红外图像获取成本的方法, 具有较高的实用价值。

**关键词:** 关联成像; 红外图像; 深度学习; 高分辨率重构

中图分类号: TP312

文献标志码: A

文章编号: 2095-2163(2024)05-0144-06

## High resolution reconstruction of near infrared association imaging based on CGAN

GUO Liang

(Security Department, Shanghai University of Engineering Science, Shanghai 201620, China)

**Abstract:** Due to complex weather conditions and low visibility at night, infrared image technology is widely used, but the cost of acquisition is high. In order to reduce the acquisition cost of infrared images, the two-dimensional Fourier single-pixel imaging method is used to obtain infrared images with low sampling rate, and then the conditional adversarial generative network is used to further reconstruct the clear infrared images. According to multiple experiments, the results show that at a sampling rate of 6.25%, infrared images close to the original clear images can be reconstructed, so the method of reducing the cost of infrared image acquisition by using associative imaging has high practical value.

**Key words:** correlated imaging; infrared images; deep learning; high-resolution reconstruction

## 0 引言

红外成像是利用探测器获得物体的红外辐射线, 背景与物体存在红外线差, 从而得到的红外图像。由于器件限制, 红外图像大多是由电偶耦合器件 (Charge Couple Device, CCD) 获取, 而由桶探测器获取红外关联像的成本更低, 在这方面, 国内外几乎很少推出相关研究成果。

近年来, 一种新颖的单像素成像技术 (Single Pixel Imaging, SPI) [1] 受到了广泛的关注, 就是通过没有空间分辨率的单像素桶式探测器重构物体信息。1995年, Pittman 等学者 [2] 通过实验实现了纠缠光的鬼成像 (Ghost Imaging, GI)。2002年, Bennink 等学者 [3] 使用经典光源开展了鬼成像实

验。Bromberg 等学者 [4] 在 2008 年通过使用空间光调制 (Spatial Light Modulation, SLM) 设备实现了计算鬼成像 (Computing Ghost Imaging, CGI)。SPI 由 GI 发展而来, 但是以上成像技术重构图像都具有明显的噪声。为此, Guo 等学者 [5] 提出了归一化鬼成像 (NGI) 来提高图像质量和抗噪性能。2008年, Duarte 等学者 [6] 提出了一种基于压缩传感的 SPI 技术, 降低了 SPI 的测量数量, 提高了成像能力。

近年来, 使用深度学习进行高分辨率图像的重构逐渐兴起, 与传统的基于插值法和基于重构法相比, 具有清晰度好、精度高等优点。2014年, Dong 等学者 [7] 首次提出卷积神经网络 (Super-Resolution Convolutional Neural Networks, SRCNN), 用于处理图像超分辨率重构; 之后又提出改进型算法 FSRCNN。

**基金项目:** 国家自然科学基金 (62275153, 62005165); 上海市产业协同创新项目 (HCXBCY-2022-006)。

**作者简介:** 郭亮 (1981-), 男, 学士, 主要研究方向: 计算机应用技术。Email: guoliang@sues.edu.cn

**收稿日期:** 2023-07-28

2016年, Shi等学者<sup>[8]</sup>提出深度卷积神经网络(Very Deep Convolutional Networks, VDSR)用来处理图像超分辨率重构。2018年, Haris等学者<sup>[9]</sup>提出深度反投影网络(Deep Back-Projection Networks, DBPN), 进一步提高了图像的超分辨率重构质量。目前, 使用深度学习进行可见光图像的超分辨率重构取得了一定的成功, 而在红外图像上的应用, 白皓等学者<sup>[10]</sup>利用深度残差网络来对红外图像进行超分辨率重构, 来提高图像的细节。与可见光图像相比, 红外图像是通过获得红外热辐射线而形成的, 分辨率差、信噪比低、视觉效果模糊、对比度低。因而之前所提到的用于可见光图像的超分辨率重构网络, 并不能直接用于红外图像的超分辨率重构中。

与其他红外图像相比, 通过傅里叶单像素成像获得的红外图像, 分辨率、对比度和视觉效果等更差, 本文主要研究条件对抗生成性网络(Conditional Generative Adversarial Networks, CGAN), 能解决红外成像分辨率低的问题。其中包括红外图像信噪比低、对比度低等问题。本文利用 NIR-Face 数据集<sup>[11]</sup>, 经过二维傅里叶单像素成像系统, 构造出低采样率后的红外图像集; 对获得的红外图像集进行预处理操作, 并将处理后的图像集与原图像集一一配对, 构造实验过程中需要的数据集; 最后, 根据红外图像特征, 采用深度学习对图像进行重构, 恢复了图像的质量。利用 CGAN 技术将桶探测器获取的低采样率图像重构为高采样率图像, 无需价格高昂的 CCD 设备, 大大降低了红外图像获取成本。

## 1 原理

### 1.1 二维傅里叶单像素成像

关联成像是一种非局域性成像技术, 近年来研究者们在该技术领域取得了一些研究成果。关联成像采用无空间分辨能力的单像素探测器, 结合空间光调制技术, 通过关联算法重构二维空间信息。关联成像技术已有近 30 年的发展历史<sup>[12]</sup>, 共经历了 3 个阶段。第一阶段为利用双光子纠缠进行成像; 第二阶段为热光关联成像, 即使用经典热光源来模拟量子纠缠从而进行关联成像; 第三阶段为计算关联成像(Computational Imaging, CI), 该技术依赖算法来测量图像中间形成过程, 而克服了传统光学和传感器硬件的限制。关联成像虽然比传统成像有更强大的抗干扰能力, 但仍然存在一定的局限性, 如成像所需的时间长、对复杂物体的还原度不高等, 因而演变出了如今的单像素成像技术。

单像素成像通过使用单像素探测器, 采集目标物体反馈的光强信息, 再结合相关算法实现目标物体的成像。其成像原理如图 1 所示。首先激光器将一束光经过扩束镜后由数字微镜器(Digital Micromirror Devices, DMD)接收, DMD 对入射光进行调制, 获得一个掩膜, 随后将掩膜信号同图像进行叠加, 经过透镜将光强信息输入到桶探测器中, 最终利用重构算法重构除场景图像。

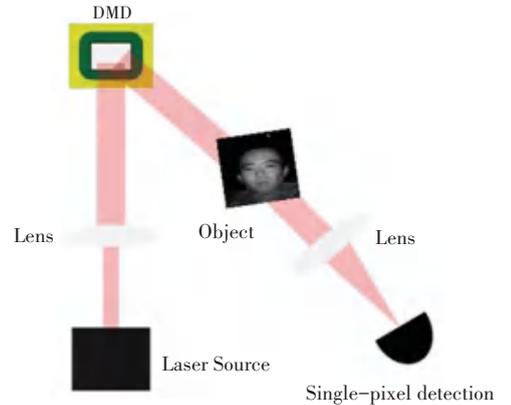


图 1 单像素成像原理图

Fig. 1 Single-pixel imaging schematic

本文采用二维傅里叶单像素成像技术<sup>[13]</sup>, 通过获取图像的傅里叶谱来对图像进行重构, 由于图像在傅里叶域的稀疏性, 可以实现高效率的成像。为获取物体图像傅里叶谱, 首先使用空间光调制器产生傅里叶基底图案, 然后对物体进行照明, 使用单像素探测器来对物体的光场强度进行探测。在物体图像的傅里叶空间进行扫描, 就可获取到图像傅里叶谱, 进行逆傅里叶变换, 实现单像素二维成像。

假设目标为反射型物体, 且物体在光照方向的反射强度为  $R(x, y)$ , 物体图像为  $I(x, y)$ , 则单像素探测器测量的光响应值可用式(1)表示:

$$D_{\varphi}(f_x, f_y) = D_n + \beta \times \iint_s R(x, y) \times P(x, y; f_x, f_y, \varphi) dx dy \quad (1)$$

其中,  $D_n$  是由探测器位置的背景照明引起的光响应值,  $\beta$  与探测器和物体的空间关系有关。为了得到对应空间频率  $(f_x, f_y)$  的傅里叶系数  $C(f_x, f_y)$ , 将空间频率  $(f_x, f_y)$  和初始相位设置为  $0, \pi/2, \pi, 3\pi/2$  四个傅里叶基模式, 分别投影到目标对象, 并记录得到的 4 个响应值  $D_1, D_2, D_3, D_4$ 。根据四步相移算法, 得到式(2):

$$C(f_x, f_y) = (D_1 - D_3) + j(D_2 - D_4) \quad (2)$$

其中,  $j$  是虚数单位。

通过得到的傅里叶系数,对图像的傅里叶谱进行逆傅里叶变换对图像进行重构,为此用到式(3):

$$I(x,y) = F^{-1}\{C(f_x, f_y)\} = F^{-1}\{(D_1 - D_3) + j(D_2 - D_4)\} \quad (3)$$

其中,  $F^{-1}$  表示傅里叶逆变换。

该方法具有高信噪比并可以减少测量次数,大大改善了单像素成像质量低、所需测量次数多的现状。

## 1.2 深度学习

经过单像素成像获得低采样率的红外图像数据集后,就可进行下一步的深度优化。在单像素成像之后,图像中包含的特征信息大部分丢失,并且出现边缘模糊问题,因此引入深度学习方法,学习模糊图像与原图之间的映射,重构出清晰图像。

GAN<sup>[14]</sup>是一种对抗生成网络模型,在需要从现有图像中生成新图像的真实场景中表现尚佳。GAN主要由生成器和判别器两个神经网络组成。生成器根据输入的图像生成新的图像,判别器则要判断生成的图像是否与真实图像一致。本文使用条件生成

式对抗网络(CGAN)<sup>[15]</sup>,是对原始GAN的一个条件性变体,生成器生成具有特定特征的样本,而不是从某个噪声分布生成一个通用样本;将特征作为输入信息,从而实现条件生成对抗式网络CGAN。本文中生成器 $G$ 生成清晰红外图像 $x$ 作为重构出来的图像 $G(x)$ ,尽可能地接近标签图像,最后利用判别器 $D$ 来判断输出结果是否与真实图像一致。

### 1.2.1 CGAN网络框架

生成器使用U-Net架构<sup>[16]</sup>,对输入图像 $x$ 进行编码,然后解码成图像 $G(x)$ 。U-Net网络使用很少的图像就能进行端到端的训练,并取得良好的效果。而这来自于完全卷积网络(FCN)<sup>[17]</sup>,基于编码解码器模型,而编码器和解码器是对称的结构。为了更好地判断图像的局部映射,判别器网络使用Patch-GAN结构,将图像划分为多个区域块,并确定每个区域块的真值和假值,最后以平均值作为最终值。网络整体框架如图2所示。

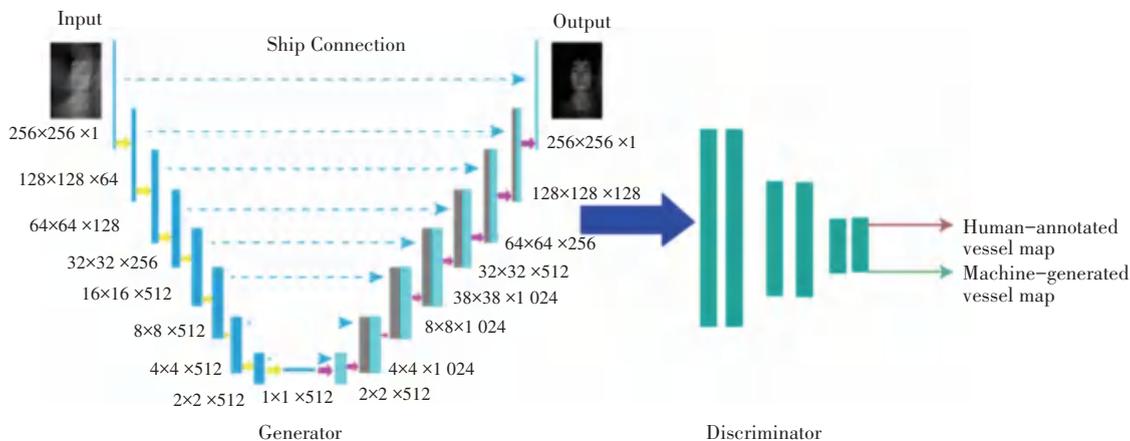


图2 CGAN网络框架

Fig. 2 CGAN network framework

### 1.2.2 损失函数

在网络训练过程中,需要使用合适的损失函数,对于条件生成对抗网络(CGAN),目标损失函数计算公式如下:

$$L_{cGAN}(G,D) = E_y[\log D(y)] + E_{x,z}[\log(1 - D(x,G(x,z)))] \quad (4)$$

其中,生成器 $G$ 不断最小化目标函数,而判别器 $D$ 则通过不断迭代来最大化目标函数。为了使输入输出图像的相似度更高,该框架中添加了 $L1$ 损失函数来计算相似性。同时,针对红外图像边缘模糊、细节缺失等特点,本文加大了相似性函数在目标损失函数中的权重。因此,该网络使用的目标损失函数为:

$$G^* = \arg \min_G \max_D L_{cGAN}(G,D) + \lambda \times L_{L1}(G) \quad (5)$$

其中,  $\lambda = 500$ 。

## 2 仿真实验

本节通过仿真实验对上述理论进行分析,实验中主要采用NIR\_Face数据集。该数据集中包含了197个人,每个人根据是否戴眼镜、不同表情、背景不同等情况拍摄了多张图片,共计3 939张图像。其中,每张人像的五官都很清晰,通过对拍摄图片进行裁剪,构成数据集。

本节分析了在不同采样率下使用CGAN重构图像的性能差异。实验过程阐述如下。首先,在Matlab 2019中通过傅里叶单像素成像,对仿真数据

集进行了降低采样率处理, 采样率为 6.25%、12.50%、25.00% 的图片各 3 939 张; 再将低采样率的图片与原图拼接成新的数据集构造训练集。神经网络的训练在 Python 3.7 中实现, 深度学习框架为 TensorFlow2.0。以上训练均在服务器上运行, 服务器配备 GPU 为 RTX2080Ti、CPU 为 i9-10900K, 性能优异。

### 2.1 评价指标

本文从主观评价与客观评价两个方面分析重构图像质量。其中, 主观评价方法主要通过观察重构图像的直观视觉效果进行图像质量评价; 客观评价方法采用峰值信噪比 (PSNR) 和结构相似度 (Structural SIMilarity Index, SSIM) 作为重构图像质量评价指标。其中, PSNR 为最大像素值与均方误差的比值, 值越大、重构图像的失真越小。其计算公式具体如下:

$$PSNR = 10 \times \lg\left(\frac{MAX_I^2}{MSE}\right) \quad (6)$$

其中, MSE 为均方根误差, MAX 为图像像素的最大值。当给定一个大小为  $m \times n$  的干净图像  $I$  和一个噪声图像  $K$ , 均方根误差公式具体如下:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|I(i, j) - K(i, j)\|^2 \quad (7)$$

结构相似度用来衡量 2 幅图像相似度, 取值范围为  $[0, 1]$ ,  $S$  值越大, 表示图像之间的相似度越高。SSIM 的计算公式具体如下:

$$SSIM(Y, \hat{Y}) = \frac{(2\mu_Y \mu_{\hat{Y}} + C)(2\sigma_{Y\hat{Y}} + C')}{(\mu_Y^2 + \mu_{\hat{Y}}^2 + C)(\sigma_Y^2 + \sigma_{\hat{Y}}^2 + C')} \quad (8)$$

其中,  $Y$  表示原始高分辨率图像;  $\hat{Y}$  表示重构图像;  $\mu$  表示图像的均值;  $\sigma$  表示图像的方差;  $\sigma_{Y\hat{Y}}$  表示 2 幅图像的协方差;  $C, C'$  表示接近 0 的正常数。

### 2.2 深度学习图像重构分析

图像模糊程度和采样率的变化成正比, 采样率越低、图像越模糊, 采样率越高、图像越接近原图。在低采样率的情况下, 图像的轮廓和细节等信息基本丢失, 因此需要使用 CGAN 神经网络作为图像重构工作的进一步优化措施。

首先, 对 NIR\_Face 数据集进行预处理。先将红外图像调整到训练需要的  $256 \times 256$  大小, 得到图像  $Tm(x)$ ; 再采用上述傅立叶方法进行降低采样率处理得到图像  $Pm(x)$ ; 由于是将可见光图像的重构算法 CGAN 引入红外图像的重构中, 所以使用 Photoshop 对图像  $Tm(x)$  和  $Pm(x)$  进行了伪彩色处理, 转换为 3 通道的彩色图像, 得到了图像  $Tm1(x)$

和  $Pm1(x)$ ; 将得到的图像对应拼接为  $Pm1(x) - Tm1(x)$  图像对。随后使用二维傅里叶单像素成像并将采样率设置为 6.25%、12.50%、25.00%。共获取 11 787 对图像, 将其中 9 000 对图像作为 CGAN 的训练集, 2 787 对图像作为测试集。预处理后的数据集称为 NIR\_Face\_E。数据集如图 3 所示。由图 3 可知, 随着采样率的升高, 图像越来越清晰。

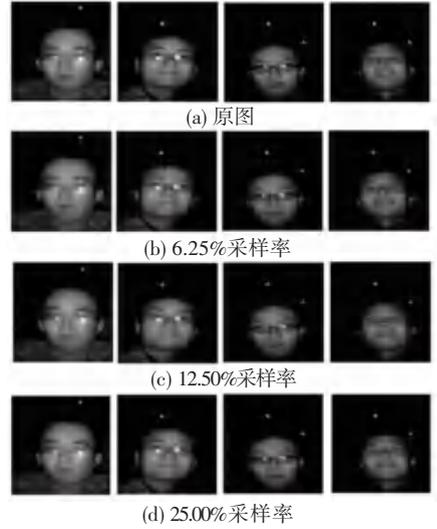


图 3 NIR\_Face\_E 训练数据集中 4 幅人脸 3 种不同采样率结果  
Fig. 3 Results of 4 faces with 3 different sampling rates in the NIR\_Face\_E train dataset

图 4 中展示了测试集图像在 6.25%、12.50% 和 25.00% 三种不同采样率下的重构结果, 客观评价指标见表 1。由表 1 可以看出, 当采样率降低为 6.25% 的时候, 重构出来的图像与原图基本接近, 并且根据表 1 得知, PSNR 和 SSIM 的值与采样率为 12.50% 与 25.00% 的时候非常接近。

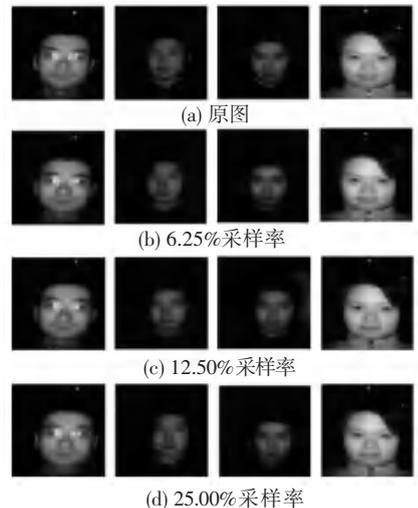


图 4 NIR\_Face\_E 测试数据集中 4 幅人脸 3 种不同采样率重构结果  
Fig. 4 Reconstruction results of 4 faces with 3 different sampling rates in the NIR\_Face\_E test dataset

表1 不同采样率下红外图像的PSNR和SSIM

Table 1 PSNR and SSIM of infrared images at different sampling rates

红外图像	6.25%		12.50%		25.00%	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
1	35.06	0.95	35.17	0.95	35.03	0.95
2	38.72	0.93	39.13	0.96	37.96	0.94
3	40.19	0.97	38.70	0.96	40.00	0.97
4	33.14	0.95	33.53	0.95	33.34	0.95

### 2.3 实验分析

为了验证仿真数据训练的可行性,本文利用了实验装置拍摄了一组红外图像,并放入由仿真数据训练好的CGAN网络中进行测试。本文使用的实验装置及其光路以激光二极管为光源,使用波长为1064 nm的激光器。当相机镜头型号为Nikon, AF-SDX55-200 mmf/4-5.6GED(68 mm×79 mm)时,所使用的桶探测器型号为POINT GRE, BFLY-PGE-50H5M(29×29×30 mm<sup>3</sup>)。捕获卡型号为M2i.2030-exp,如图5所示。



图5 实验装置图

Fig. 5 Experimental setup diagram

利用实验设备拍摄了一组小玩偶的红外图像,然后将拍摄的图像调整为采样率6.25%。激光器首先通过DMD获取调制光场,并于目标玩偶图像进行叠加,光路经过反射后由桶探测器接收,以获得光信号。将获取的低采样放入训练好的CGAN网络中进行图像重构测试。测试结果如图6所示。



(a) 采样率6.25%输入图像 (b) 真实图像 (c) CGAN重建图像

图6 测试结果图

Fig. 6 Test results graph

从图6中可以看出,经过低采样率获取的图像模糊,且丢失较多信息,仅能保留原图大致特点。而

将低采样率图像经过CGAN网络进行重构则可复原出较多信息,接近原图。重构量化指标见表2。

表2 重构前后图像对比

Table 2 Comparison of images before and after reconstruction

图像	PSNR	SSIM
采样率6.25%输入图像	10.091 1	0.737 2
CGAN重构后图像	13.830 7	0.839 7

由表2可以看出相较于低采样率下的输入图像重构出来图像的PSNR值增加了36.72%,SSIM值增加了13.90%,有较大提升,因而仿真数据训练出来的网络是可行的。

### 3 结束语

本文针对红外图像获取成本较高的问题提出基于CGAN网络对低采样率下的模糊红外关联成像图像进行高分辨率重构。实验结果表明,该方法可以显著提高低采样率下的红外图像质量。

虽然CGAN可以重建高清图像,但是当网络重建图像类型与网络所用训练图像类型不同时往往难以获得足够的细节。为了克服这一问题,下一步的研究是利用迁移学习让网络获得不同类型图像的特征以提高细节的重建精度。

### 参考文献

- [1] CAI Yangjian, ZHU Shijun. Ghost imaging with incoherent and partially coherent light radiation[J]. Physical Review E, 2005, 71(5): 056607.
- [2] PITTMAN T B, SHIH Y H, STREKALOV D V, et al. Optical imaging by means of two-photon quantum entanglement[J]. Physical Review A, 1995, 52(5): R3429.
- [3] BENNINK R S, BENTLEY S J, BOYD R W, et al. Two-photon coincidence imaging with a classical source[J]. Physical Review Letters, 2002, 89(11): 113601.
- [4] BROMBERG Y, KATZ O, SILBERBERG Y. Ghost imaging with a single detector[J]. Physical Review A, 2008, 79(5): 1744 - 1747.
- [5] GUO Shuxu, ZHANG Chi, CAO Junsheng, et al. Object reconstruction by compressive sensing based normalized ghost imaging[J]. Optics and Precision Engineering, 2015, 23(1):

- 288–294.
- [6] DUARTE M F, DAVENPORT M A, TAKHAR D, et al. Single-pixel imaging via compressive sampling [J]. *IEEE Signal Processing Magazine*, 2008, 25(2): 83–91.
- [7] DONG Chao, CHEN C L, HE Kaiming, et al. Learning a deep convolutional network for image super-resolution [C]// *Proceedings of the 2014 European Conference on Computer Vision*. Berlin/Heidelberg: Springer, 2014, 8692: 184–199.
- [8] SHI Wenzhe, CABALLERO J, HUSZAR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network [C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, USA: IEEE, 2016: 1874–1883.
- [9] HARIS M, SHAKHNAROVICH G, UKITA N. Deep back-projection networks for super-resolution [C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA: IEEE, 2018: 1664–1673.
- [10] 白皓, 白廷柱. 基于深度残差神经网络的红外图像超分辨率重构算法 [J]. *红外技术*, 2024, 46(2): 176–182.
- [11] LI S Z, CHU Rufeng, LIAO Shengcai, et al. Illumination invariant face recognition using near-infrared images [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence (Special issue on Biometrics: Progress and Directions)*, 2007, 29(4): 627–639.
- [12] 赵剡, 李高亮, 杨照华. 量子关联成像技术发展 [J]. *航空兵器*, 2017(5): 3–10.
- [13] ZHANG Leihong, BIAN Zhixiang, YE Huahong, et al. Restoration of Single pixel imaging in atmospheric turbulence by Fourier filter and CGAN [J]. *Applied Physics B – Lasers and Optics*, 2021, 127(3): 45.
- [14] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative Adversarial Networks [C]// *Advances in Neural Information Processing Systems*. Montreal, Canada: NIPS Foundation, 2014, 3: 2672–2680.
- [15] UEHARA M, SATO I, SUZUKI M, et al. Generative adversarial nets from a density ratio estimation perspective [J]. *arXiv preprint arXiv:1610.02920*, 2016.
- [16] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation [J]. *arXiv preprint arXiv:1505.04597*, 2015.
- [17] LONG J, SHELHAMER E, DARRELL T. Fully Convolutional Networks for Semantic Segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 39(4): 640–651.